

What is claimed is

1. A method of organizing information into a plurality of classes or clusters with a user-configurable information clustering system comprising:
  - a) grouping units of information into clusters based on similarities to create a cluster structure; and
  - b) personalizing said cluster structure according to user knowledge and preferences.
2. The method according to claim 1 wherein said grouping units of information into clusters is carried out automatically to create a machine-generated cluster structure.
3. The method according to claim 1 wherein said personalizing comprises creating at least one new information cluster.
4. The method according to claim 3 wherein said personalizing further comprises labeling each information cluster.
5. The method according to claim 4 wherein said personalizing further comprises merging information clusters.
6. The method according to claim 5 wherein said personalizing further comprises splitting at least one information cluster.
7. The method according to claim 6 wherein said personalizing further comprises storing said cluster structure in a knowledge base.
8. The method according to claim 1 wherein said personalizing comprises labeling each information cluster.
9. The method according to claim 1 wherein said personalizing comprises merging information clusters.

10. The method according to claim 1 wherein said personalizing comprises splitting at least one information cluster.
11. The method according to claim 1 wherein said personalizing comprises storing said cluster structure in a knowledge base.
12. The method according to claim 1 wherein said information comprises text, image, audio, video or any combination thereof.
13. The method according to claim 1 wherein said user-configurable information clustering system comprises an adaptive resonance associative map.
14. The method according to claim 1 wherein said user-configurable information clustering system incorporates user knowledge and preferences for information clustering.
15. The method according to claim 1 wherein said user-configurable information clustering system further comprises a user interface.
16. The method according to claim 1 wherein each of said units of information is represented by an information vector.
17. The method according to claim 1 wherein a user-preferred information grouping is represented by a preference vector.
18. The method according to claim 1 wherein said units of information are grouped into classes or clusters based on a similarity function.
19. The method according to claim 18 wherein said classes or clusters have a coarseness which is controlled by a baseline vigilance parameter.

20. The method according to claim 1 further comprising indication by a user of a preference for a lower baseline vigilance parameter by selecting at least one unit of information from each of at least two clusters wherein the selected units of information are deemed by the user to be similar to each other.
21. The method according to claim 1 further comprising indication by a user of a preference for a higher baseline vigilance parameter by selecting at least two units of information in a cluster, wherein said units of information are deemed by the user to be dissimilar to each other.
22. The method according to claim 1 further comprising retrieving said cluster structure to initialize said user-configurable information clustering system prior to clustering new information.
23. A method of building information classification systems comprising:
- a) grouping information vectors into clusters based on similarities;
  - b) assigning a category label to each cluster;
  - c) merging information clusters;
  - d) splitting information clusters; and
  - e) storing labeled clusters as the defined categories of a classification system.
24. The method according to claim 23 wherein said information comprises text, image, audio, video or any combination thereof.
25. A method for new information detection and trend analysis with a user-configurable information clustering system comprising:
- a) grouping information vectors into clusters based on similarities;
  - b) creating a labeled cluster structure by assigning a category label to each cluster, merging information clusters and splitting information clusters according to a user's preferences;

- c) storing said labeled cluster structure, wherein said cluster structure defines the user's knowledge;
  - d) retrieving said cluster structure;
  - e) initializing the information clustering system using said retrieved cluster structure; and
  - f) analyzing new clusters, wherein said clusters are grouped according to the user's preferences.
26. The method according to claim 25 wherein said information comprises text, image, audio, video or any combination thereof.
27. A user-configurable information clustering system comprising:
- a) an information clustering engine for clustering units of information based on similarities to create a cluster structure;
  - b) a personalization module for personalizing said cluster structure according to user knowledge and preferences;
  - c) a user interface; and
  - d) a knowledge base for storing said cluster structure.
28. The system according to claim 27 wherein said information clustering engine automatically clusters information to create a machine-generated cluster structure.
29. The system according to claim 27 wherein said personalization module comprises means for creating at least one new information cluster.
30. The system according to claim 29 wherein said personalization module further comprises means for labeling each information cluster.
31. The system according to claim 30 wherein said personalization module further comprises means for merging information clusters.

32. The system according to claim 31 wherein said personalization module further comprises means for splitting at least one information cluster.
33. The system according to claim 32 wherein said personalization module further comprises means for storing the cluster structure in said knowledge base.
34. The system according to claim 33 wherein said personalization module further comprises means for retrieving the cluster structure from said knowledge base.
35. The system according to claim 27 wherein said personalization module comprises means for labeling each information cluster.
36. The system according to claim 27 wherein said personalization module comprises means for merging information clusters.
37. The system according to claim 27 wherein said personalization module comprises means for splitting at least one information cluster.
38. The system according to claim 27 wherein said personalization module comprises means for storing the cluster structure in said knowledge base.
39. The system according to claim 27 wherein said personalization module comprises means for retrieving the cluster structure from said knowledge base.
40. The system according to claim 27 wherein said information comprises text, image, audio, video or any combination thereof.
41. The system according to claim 27 wherein user knowledge and preferences are incorporated in information clustering.

42. The system according to claim 27 wherein said information clustering engine comprises an adaptive resonance associative map.
43. The system according to claim 27 wherein said user interface permits graphical visualization of said information clusters.
44. The system according to claim 27 wherein each of said units of information is represented by an information vector.
45. The system according to claim 27 wherein a user-preferred information grouping is represented by a preference vector.
46. The system according to claim 27 wherein said units of information are grouped into classes or clusters based on a similarity function.
47. The system according to claim 46 wherein said classes or clusters have a coarseness which is controlled by a baseline vigilance parameter.
48. The system according to claim 27 wherein said personalization module permits indication by a user of a preference for a lower baseline vigilance parameter by selecting at least one unit of information from each of at least two clusters wherein said selected units of information are deemed by the user to be similar to each other.
49. The system according to claim 27 wherein said personalization module permits indication by a user of a preference for a higher baseline vigilance parameter by selecting at least two units of information in a cluster, wherein said units of information are deemed by the user to be dissimilar to each other.
50. An information classification system comprising:
  - a) means for grouping information vectors based on similarities;
  - b) means for creating a plurality of information clusters;

- c) means for labeling each information cluster;
- d) means for merging information clusters;
- e) means for splitting at least one information cluster; and
- f) means for storing labeled clusters as the defined categories of a classification system.

51. The system according to claim 50 wherein the information comprises text, image, audio, video or any combination thereof.

52. A system for new information detection and trend analysis comprising:

- a) means for grouping information vectors into clusters based on similarities;
- b) means for creating a plurality of information clusters;
- c) means for assigning a category label to a cluster;
- d) means for merging information clusters;
- e) means for splitting at least one information cluster;
- f) means for storing labeled clusters wherein said clusters define a user's knowledge of said information.
- g) means for retrieving said clusters to permit analysis of new clusters, wherein said new clusters are grouped according to the user's preferences.

53. The system according to claim 52 wherein the information comprises text, image, audio, video or any combination thereof.

TO 4099 "T 224360